

Cassandra Crossing/Intelligenza Artificiale: Esorcismi

(532)—Cosa fare quando ci si accorge di aver creato un demone invece di un angelo?

Cassandra Crossing/Intelligenza Artificiale: Esorcismi



Figure 1:

(532)—Cosa fare quando ci si accorge di aver creato un demone invece di un angelo?

10 febbraio 2023—Avete avuto modo di giocherellare con la falsa “Intelligenza Artificiale” più popolare, la mitica ChatGPT?

Fatelo; e mentre voi porrete divertenti domande a GPT-3, il software che sta dietro ChatGPT, dei veri apprendisti stregoni saranno al lavoro, tentando, dopo averlo evocato, di domare il demone e di farlo rimanere nel pentacolo.

Troppo dark e poco comprensibile, anche per molti dei 24 indefessi lettori? Allora fermiamoci, perché servono alcune premesse.

Iniziamo fornendo, a chi non fosse al corrente dell’esistenza di un’intelligenza artificiale allo stato dell’arte che chiunque può interrogare a piacimento, [direttamente il link a cui trovarla](#), ma invitandoli prima a leggere questa raffica di esternazioni di Cassandra, se già non l’avessero fatto:

“Ragionare usando i dati”, “L’IA oltre i suoi limiti”, “IA: il sortilegio”, “IA: il Potere” ed “IA: il Programmatore pericoloso”.

Inoltre, a vantaggio di chi non fosse un cultore di esoterismo o di film horror, ma anche come ripasso per chi qualche nozione ne possieda, ricordiamo che un mortale può evocare un’entità ultraterrena eseguendo un esorcismo, di solito tramite una formula magica che deve essere pronunciata ad alta voce.

Ma poiché l’entità potrebbe essere molto maldisposta verso chi l’ha evocata, in parole povere potrebbe non essere angelica ma demoniaca, è necessario proteggersi, tracciando sul pavimento un disegno detto “pentacolo”, che il demone evocato non è in grado di oltrepassare, imprigionandolo.

E’ anche possibile utilizzare il pentacolo al contrario, ponendosi al suo interno per proteggersi mentre si pronuncia l’esorcismo, lasciando però in questo caso il demone libero di scorrazzare nel mondo.

E torniamo a noi. In realtà ChatGPT altro non è che una enorme e ben congegnata campagna pubblicitaria, costruita per propagandare la falsa Intelligenza Artificiale di OpenAI al grande pubblico.

E’ stata progettata evidentemente da dei tecnologi, per loro stessa natura incapaci di pensare alla conseguenze di una loro opera prima di lanciarla nel mondo, anche perché molto incoraggiati da chi vuol farci i soldi.

Infatti, poiché i modelli linguistici avanzati come GPT-3 vengono nutriti con quantità enormi di informazioni, provenienti perlopiù da Internet, il bene ed il male che permeano la Rete delle reti, che è solo un’immagine dell’umanità, vi vengono riversati tal quali.

E quindi, interagendo con le false Intelligenze Artificiali, potremmo trovarci di fronte non ad un cucciolo puffy ma ad un lupo feroce. Una semplice domanda a ChatGPT potrebbe evocare lo spirito di uno dei demoni che albergano nelle menti degli uomini.

Ora, se ChatGPT iniziasse a parlare come Chtulu in persona forse non succederebbe niente, le nuove tecnologie possono anche mostrare problemi. Anche se desse suggerimenti dual-use, come le dosi di arsenico necessarie ad eliminare sia topi che uomini, forse andrebbe ancora bene.

Al massimo verrebbe citato “Lo strano caso del dottor Jekyll e mister Hyde”

Ma ... e se ChatGPT dicesse qualcosa di razzista, maschilista, negazionista, omofobico, anti-meetoo, anti-green? Se scrivesse una nuova storia, quella de “Lo strano caso del dottor Elon e mister Musk”? Se insomma violasse uno, anzi tutti i tabù del politically correct più deteriore? Quelli per cui vengono talvolta distrutte persone vere che dicono quello che pensano?

E’ già successo, ad altre false Intelligenze Artificiali, già tante volte.

E sì, anche ChatGPT, appena gli hanno fatto aprire bocca, ne ha dette ...

Ecco che gli azionisti di OpenAI e delle altre startup di false Intelligenze Artificiali hanno scoperto di aver evocato il loro demone, la loro nemesi. Hanno scoperto di aver evocato un’entità che invece di arricchirli di buona pubblicità potrebbe annientarli in uno shitstorm a banda larghissima.

Ed allora hanno preso provvedimenti. Non hanno certo ucciso la gallina da cui si aspettano le uova d’oro. No, hanno semplicemente costruito una moderna versione del pentacolo, dentro cui evocare GPT-3.

Sempre più persone vengono impiegate per “ascoltare” le domande poste a ChatGPT e costruire filtri e vincoli che impediscono al demone, alla falsa Intelligenza Artificiale, di ascoltare certe domande o di fornire certe risposte. Una specie di moderno pentacolo, da cui solo risposte politically correct possono uscire.

Ma all’interno del pentacolo resta comunque il demone, la falsa Intelligenza Artificiale che, pur incapace di pronunciare le risposte più infelici, continuerà a rispondere comunque con i suoi limiti.

Fornirà risposte prodotte da un demone che non sa niente e non conosce niente, risposte piene delle insignificanti ceneri di tutte le informazioni che le sono state riversate dentro e incenerite; da queste ceneri è sorto un mostro di statistica cancerosa, che ammalia la gente perché sa usare le parole giuste, ma può comporre solo per produrre frasi vuote, lasciando che poi ciascuno le riempia con i significati che vuole. Non demonio tecnologico quindi, ma solo una più moderna ma sempre insignificante Eliza.

Alla fine di questa cervellotica ed un po’ ermetica esternazione, resta solo una domanda. In fondo il problema non è ChatGPT, ma dove noi ci troviamo rispetto a lui; siamo dentro o fuori dal pentacolo?

Nota per che riceve gli articoli via mail. Medium.com modifica automaticamente i link contenuti negli articoli quando li invia per mail, rendendoli traccianti. ***La cosa disgusta Cassandra***, che se ne è accorta solo di recente grazie ad una provvidenziale segnalazione. Se ciò superasse il vostro limite di indignazione, ed in attesa che Cassandra trovi una soluzione od un’alternativa, potete fruire dell’articolo direttamente sul sito.

[Scrivere a Cassandra—Twitter—Mastodon](#)
[Videorubrica “Quattro chiacchiere con Cassandra”](#)
[Lo Slog \(Static Blog\) di Cassandra](#)
[L’archivio di Cassandra: scuola, formazione e pensiero](#)

Licenza d’utilizzo: i contenuti di questo articolo, dove non diversamente indicato, sono sotto licenza Creative Commons Attribuzione—Condividi allo stesso modo 4.0 Internazionale (CC BY-SA 4.0), tutte le informazioni di utilizzo del materiale sono disponibili a [questo link](#).

By [Marco A. L. Calamari](#) on [February 12, 2023](#).

[Canonical link](#)

Exported from [Medium](#) on August 27, 2025.